

# 低消費電力で高性能を実現した画像認識プロセッサ Visconti™ 2シリーズ

Visconti™ 2 Series Image Recognition Processors with High Performance and Low Power Consumption

田邊 靖貴      荒木田 英穂      伴野 守保

■ TANABE Yasuki      ■ ARAKIDA Hideho      ■ BANNO Moriyasu

画像認識技術の様々な製品への利用が急速に拡大している。東芝は、高い画像認識処理性能を、低消費電力で実現する画像認識プロセッサ“Visconti™ 2”(TMPV7506XBG, TMPV7504XBG)を開発した。

TMPV7506XBGでは、4台のカメラ入力に対応し、4個のメディアプロセッシングエンジン(MPE)と6個の画像処理アクセラレータを搭載する。これにより、五つの画像認識と一つの画像処理アプリケーションを同時に実行した例でも、平均約0.74Wという低消費電力を実現した。またハンドジェスチャ処理においては、1GHz動作の汎用組込み機器向けプロセッサを用いた場合と比較して、約6~10倍の性能向上を確認した。

With the advancement of image recognition technologies in recent years, their application to various types of products is rapidly spreading. Toshiba has developed two new devices in the Visconti™ 2 series of image recognition processors, the TMPV7506XBG and TMPV7504XBG, offering both high performance and low power consumption.

The TMPV7506XBG has four video inputs, four media processing engines (MPEs), and six image processing accelerators. It achieves an average power consumption of about 0.74 W during the simultaneous execution of one image processing and five image recognition applications. In the case of hand gesture recognition applications, it also achieves an improvement in performance of about six to 10 times compared with a general-purpose processor operated at 1 GHz in an embedded system.

## 1 まえがき

画像認識技術を適用し、安全、安心、及び利便性を向上させるシステムや製品によるスマートライフの実現が進んできている。自動車産業では、図1に示すような前方衝突検出などの先進運転支援システム<sup>(1)</sup>が、民生機器では、ジェスチャコントロール<sup>(2)</sup>や、監視カメラにおける侵入者検知などが、その適用例として挙げられる。

東芝は、ノートPC(パソコン)のQosmio™へのハンドジェスチャ技術の搭載や、画像認識プロセッサ“Visconti™”の開発などをはじめとして、様々な画像認識に関する技術開発を進めてきた。近年、このような画像認識技術の実用化が加速しており、更なる成長が期待されている。そこで、従来よりも高度な画像認識技術を採用した製品開発を見据え、Visconti™の後継の開発を推進した。

画像認識技術を製品に適用するには、大量の演算処理を行う必要がある。例えば、先進運転支援システムの一つのアプリケーションである歩行者検出を行う場合、画像の中から人と思われる領域を抽出し、抽出された各領域について、人であるか否かを識別する処理を行う。1枚の画像には候補となる領域が多数あり、更に動画像に対してリアルタイムで歩行者を検出して追跡する必要がある。また、前進時には歩行者検出だけでなく障害物や標識も検出するなど、複数のアプリケーシ

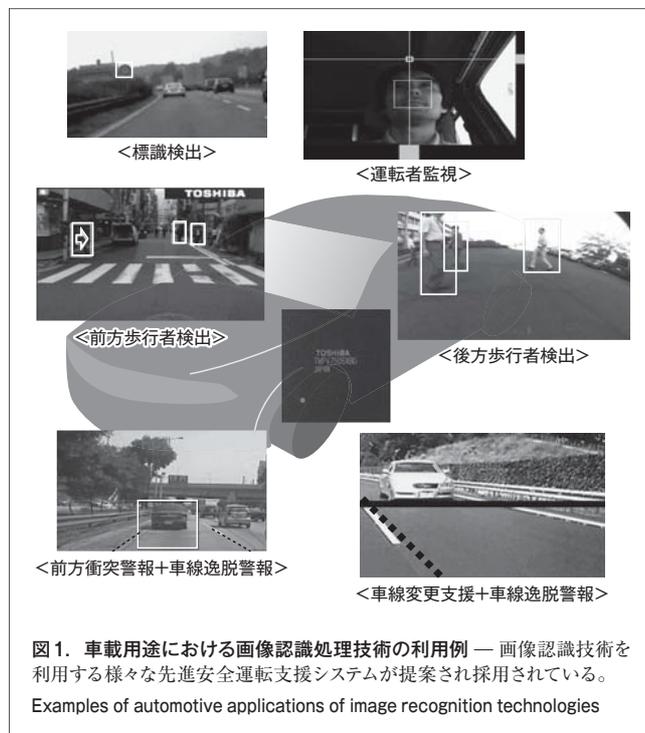


図1. 車載用途における画像認識処理技術の利用例 — 画像認識技術を利用する様々な先進安全運転支援システムが提案され採用されている。  
Examples of automotive applications of image recognition technologies

ンを同時に実行することが要求される。このように、高度な画像認識を実現するには、高い処理性能が必要となる。

その一方で、低消費電力を実現することも求められる。例

例えば車載用途では、エンジン停止時に車載バッテリーでも駆動可能な低消費電力が求められる。また、システムの信頼性を向上させるために、機械構造を持つ冷却ファンを使わないで機器が設計できることが求められる。これらの観点からも消費電力を抑えることが重要である。

そこで、膨大な演算処理を低消費電力で実現する画像認識プロセッサ“Visconti™ 2”<sup>(3)</sup> (TMPV7506XBG及びTMPV7504XBG)を開発した。ここでは、Visconti™ 2のプロセッサアーキテクチャの概要と、これを利用したアプリケーションの実装例について述べる。

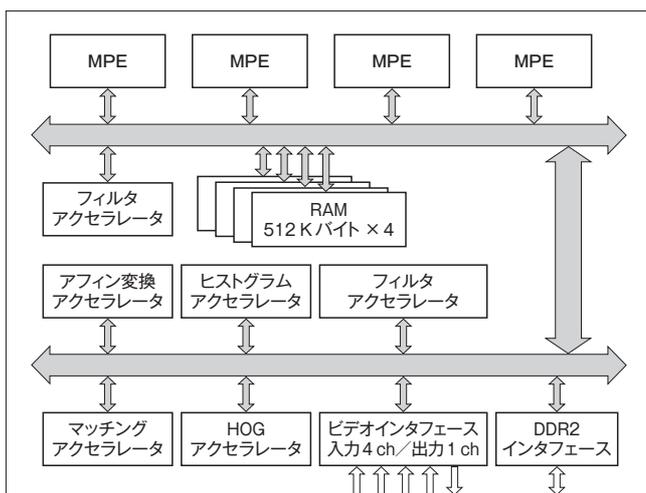
## 2 Visconti™ 2の構成

TMPV7506XBGのアーキテクチャを図2に示す。

Visconti™ 2は、266 MHz動作のメディア処理向けプロセッサ(メディアプロセッシングエンジン:MPE)を4個搭載する。これにより、アプリケーション処理の並列実行や複数アプリケーションの分散実行が可能になる。

またVisconti™ 2には、画像認識処理で頻繁に利用される処理タスクや、多くの演算量を必要とする処理タスクを実行する、後述の5種類のアクセラレータを搭載した。これにより、画像処理中の各種タスクを効率的に実行し、高い演算処理性能と低消費電力の両立を実現する。

- (1) アフィン変換アクセラレータ 鳥観表示など視点変換をした画像を生成する、広角・魚眼レンズで撮影した歪み(ひずみ)のある画像を補正する、といった画像の各種変形処理を行う。



ch: チャネル

図2. Visconti™ 2 (TMPV7506XBG)のアーキテクチャー 画像認識処理を効率よく実行するために、4個のメディア処理向けプロセッサコアとともに、最大構成で5種類、計6個のアクセラレータを搭載する。

Architecture of TMPV7506XBG image recognition processor in Visconti™ 2 series

- (2) ヒストグラムアクセラレータ 画像処理で頻繁に利用される特徴量の一つであるヒストグラムを生成する。
- (3) フィルタアクセラレータ 画像処理で頻繁に利用される輪郭画像生成や輝度勾配画像生成などのフィルタ処理を行う。複数の画素データを並列処理するために、10個の演算器を1セットとし、これを64セット並べた構成を採り、1クロックサイクル当たり最大で640の演算を実行できるようにしている。また、フィルタ処理を並列実行できるように、このアクセラレータを2個搭載している。これにより効率よくフィルタ処理を実行できるようにした。
- (4) マッチングアクセラレータ 2台のカメラを利用したステレオ視で画像中の物体までの距離を推定することにより、認識候補領域の絞込みを行える。これにより、認識処理の誤検出率を大幅に低減できる<sup>(4)</sup>。距離の推定には、ステレオ画像から対応する点を探索する、マッチング処理を行う。しかし、このマッチング処理は膨大な演算を必要とする。マッチングアクセラレータは、この処理から並列性を抽出するアプローチを用いて、1クロックサイクル当たり最大13点の対応点の探索処理を行う。これにより、従来のVisconti™ 2を利用したマッチング処理<sup>(5)</sup>と比較し、25倍以上の処理性能を実現した。
- (5) HOGアクセラレータ (TMPV7504XBGには非搭載)

HOG (Histogram of Oriented Gradients) 特徴量及びCoHOG (Co-occurrence Histogram of Oriented Gradients) 特徴量の演算と、識別処理を実行する。CoHOG特徴量を利用する識別処理は人物検出をはじめ、様々なオブジェクトに対し高い識別精度(認識対象か否かの判別)を実現する<sup>(6)</sup>。しかし、その演算にはメモリを頻繁に参照する必要がある。HOGアクセラレータでは演算に利用するデータをアクセラレータ内で効率的に再利用しつつ、演算に必要なデータを効率的に演算器に供給する構成を採る。これにより、組込み機器用の量産プロセッサとして世界で初めて<sup>(注1)</sup>CoHOG特徴量を用いた認識処理の実時間実行を実現した。

Visconti™ 2に搭載したMPEとアクセラレータの演算性能を引き出すには、各モジュールが効率よくデータをアクセスできる必要がある。そこで、Visconti™ 2では、外部DDR2 (Double Data Rate 2) SDRAM (Synchronous DRAM)に加え、512 KバイトのオンチップRAMを4バンク搭載した。これらのDDR2インタフェースやオンチップRAMは、クロスバー構造のオンチップバスにそれぞれ接続されており、MPEやアクセラレータから並列にアクセスできるようにした。

ビデオ入力インタフェースとしては、TMPV7506XBGでは4チャンネル、TMPV7504XBGでは2チャンネルを設けている。

(注1) 2011年10月時点、当社調べ。

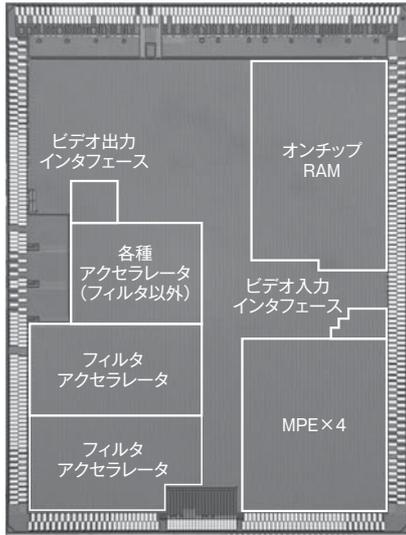


図3. TMPV7506XBGのチップ— 40 nm CMOSプロセスを採用し、44.54 mm<sup>2</sup>のチップサイズに集積している。  
Chip of TMPV7506XBG

表1. 各モジュールごとの性能及び消費電力

Performance and power consumption of each module

モジュール	周波数 (MHz)	性能値 (GOPs)	消費電力 (mW)	消費電力当たり性能 (GOPs/W)
MPE × 4	266	46.8	120.6	388.1
アフィン変換アクセラレータ	266	10.6	108.2	98.4
ヒストグラムアクセラレータ	266	5.3	42.0	126.7
フィルタアクセラレータ × 2	180	230.4	293.1	786.2
マッチングアクセラレータ	266	145.2	54.6	2,659.3
HOGアクセラレータ	266	25.5	130.2	196.1
合計		463.9	748.6	619.7

\*電源電圧1.1 V, 周囲温度25 °Cの条件で評価

これにより、自動車を上方から見たような画像を合成する画像処理や、前方衝突検出と車両左右後方の死角監視を同時に実行する画像認識といった、複数のカメラ入力を利用する処理を実現できる。

Visconti<sub>TM</sub> 2 (TMPV7506XBG) のチップを図3に示す。Visconti<sub>TM</sub> 2は40 nm CMOSプロセスを採用し、チップサイズは44.54 mm<sup>2</sup>である。

モジュールごとの最大動作周波数、性能値 (GOPs (Giga Operations per Second: 毎秒実行可能な演算数)), 及び最大消費電力の評価結果を表1に示す。Visconti<sub>TM</sub> 2 (TMPV7506XBG) のピーク性能は、4個のMPEと6個のアクセラレータの合計で約464 GOPsである。これは従来のVisconti<sub>TM</sub>の10倍以上の性能に相当する。また、消費電力当たりの性能 (GOPs/W) としても、約620 GOPs/Wという高い効率を実現した。

### 3 先進運転支援システムへの応用

自動車向けの先進運転支援システムでは、複数のアプリケーションを同時に実行するユースケースが想定される。そこでVisconti<sub>TM</sub> 2が、認識処理を含めた複数のアプリケーションを同時に実行するために、十分な性能を持つ一例を図4に示す (TMPV7506XBGで実行)。この例では、後方カメラによる車線変更支援及び車線逸脱警報と、車内カメラによる運転者監視を行う三つの画像認識、及び魚眼レンズを用いたカメラからの入力画像の歪み補正と視点変換を行う画像処理の合計4種類のアプリケーションを、同時に、それぞれ毎秒30フレームのビデオ入力に対してリアルタイムで処理している。このようにVisconti<sub>TM</sub> 2は、複数の画像認識と画像処理アプリケーションを実行するのに十分な性能を持つ。

また、各アクセラレータで処理を効率的に実行することにより、低消費電力を実現している。前述の処理に前方カメラによる前方衝突警報と車線逸脱警報を加えた、全体で一つの画像処理と、五つの画像認識アプリケーションをTMPV7506XBG上で同時に実行した例では、平均で約0.74 Wという低消費電力を実現した。

図4. TMPV7506XBG上で同時に実行したアプリケーション— 運転者監視は検出した領域を矩形で、車線逸脱警報は検出した車線を線で、車線変更支援は近づく車両を線で表示している。また前方衝突警報では、検出した前方車両を矩形で表示している。

Applications simultaneously executed on TMPV7506XBG

### 4 ジェスチャ認識処理への応用

画像認識処理は、民生機器でも利用範囲が広がっている。例えば、手の動きで機器を操作できるハンドジェスチャ ユーザーインターフェース (HGUI: Hand Gesture User Interface)

技術を搭載した機器の実用化が、一例として挙げられる。当社でも、HGUIを搭載したPCを製品化している<sup>(2)</sup>。このHGUI処理を、Visconti™2の高性能なアクセラレータを最大限に活用できるようにアルゴリズムと設計を見直した(図5)。

今回開発したHGUI処理は、手の検出を行う処理(手検出モード)と、ジェスチャを検出するモード(ジェスチャ検出モード)の二つのモードから成る。どちらのモードでも、まず多数の縮小画像を生成する処理(ピラミッド画像生成)を行う。そして手検出モードでは、この縮小画像に対し、CoHOG特徴量を用いた候補領域の生成と、手の検出処理を行う。ジェスチャ検出モードでは、前フレームで手を検出した領域の周辺領域に対して、CoHOG特徴量により、複数の手形状の検出(ジェスチャ検出)を行う。

この処理に要した時間を、1GHz動作の汎用組み込み機器向けのプロセッサで実行した場合と比較した(表2)。処理ごとの実行時間を比較すると、ピラミッド画像生成処理ではアフィン変換アクセラレータにより約15倍、ジェスチャ検出処理ではHOGアクセラレータにより約63倍の性能向上を実現した。そして処理全体では、約6~10倍という大幅な性能向上を達成した。

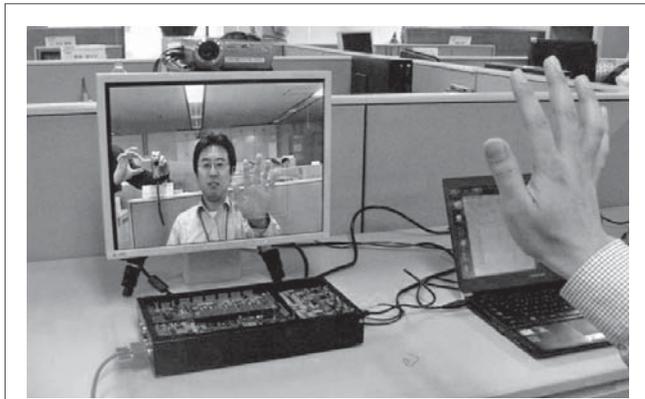


図5. ハンドジェスチャの検出 — 手検出、及びジェスチャ検出処理をVisconti™2リファレンスボード上に実装し、認識処理速度が大幅に向上することを確認した。

Hand gesture user interface implemented on TMPV7506XBG

表2. ハンドジェスチャ検出処理の実行時間の比較

Comparison of execution times of TMPV7506XBG and general-purpose processor

処理内容	実行時間 (ms)	
	Visconti™2 (最大266 MHz)	汎用プロセッサ (1 GHz)
ピラミッド画像生成	6.5	92.3
手検出処理(手検出モード)	32.4	434.5
ジェスチャ検出処理(ジェスチャ検出モード)	1.1	69.5
手検出モード全体	54.3	540.9
ジェスチャ検出モード全体	29.0	176.5

## 5 あとがき

Visconti™2のアーキテクチャとその応用例について述べた。Visconti™2は4個のMPEと画像認識に特化した6個のアクセラレータ群を搭載する。これにより、低消費電力で複数の画像認識アプリケーションを同時に実行できる性能を実現した。

画像認識技術を利用したシステムや機器による安全、安心、及び利便性の提供は、スマートライフの実現に必要不可欠と考えられる。今後Visconti™2が、これらのシステム及び機器の開発をけん引するプロセッサとして活躍することが期待される。

## 文献

- (1) 尾崎信之 他. 安全運転支援システム. 東芝レビュー. 66, 2, 2011, p.13-16.
- (2) 池 司 他. 自然な手振りによるハンドジェスチャ ユーザーインタフェース. 東芝レビュー. 67, 6, 2012, p.36-39.
- (3) Tanabe, Y. et al. "A 464GOPS 620GOPS/W heterogeneous multi-core SoC for image-recognition applications". IEEE International Solid-State Circuits Conference (ISSCC) Dig. Tech. Papers. San Francisco, CA, USA, 2012-02, p.222-223.
- (4) 服部 寛. 動的な背景の中から人物だけを検出する画像認識技術. 東芝レビュー. 65, 6, 2010, p.60-61.
- (5) 服部 寛. 車載向けステレオ画像処理技術. 東芝レビュー. 63, 5, 2008, p.48-51.
- (6) Watanabe, T. et al. "Co-occurrence Histograms of Oriented Gradients for Pedestrian Detection". Proceedings of the 3rd Pacific Rim Symposium on Advances in Image and Video Technology. Tokyo, 2009-01, National Institute of Informatics (NII) et al. p.37-47.



田邊 靖貴 TANABE Yasuki, Ph.D.

セミコンダクター & ストレージ社 半導体研究開発センター デジタルメディアSoC技術開発部主務、博士(工学)。画像処理を行うプロセッサ及びアクセラレータの研究・開発に従事。  
Center for Semiconductor Research & Development



荒木田 英穂 ARAKIDA Hideho

セミコンダクター & ストレージ社 半導体研究開発センター デジタルメディアSoC技術開発部主査。画像処理を行うプロセッサ向けファームウェアの開発に従事。  
Center for Semiconductor Research & Development



伴野 守保 BANNO Moriyasu

セミコンダクター & ストレージ社 アナログ・イメージングIC事業部 車載LSI技術部参事。プロセッサコア及びSoCの企画・開発に従事。  
Analog & Imaging IC Div.